MONOPHYLY OF THE ASTERIDAE AND IDENTIFICATION OF THEIR MAJOR LINEAGES INFERRED FROM DNA SEQUENCES OF rbcL[1]

Richard G. Olmstead,[2,3]
Helen J. Michaels,[4] Kathy M. Scott,[2]
and Jeffrey D. Palmer[2]

ABSTRACT

A parsimony analysis of 57 angiosperm rbcL sequences was conducted to test the monophyly of the Asteridae and to identify major lineages within the Asteridae. Three major clades, the Caryophyllidae, the Rosidae plus Dilleniidae, and the Asteridae sensu lato, emerge from an unresolved radiation in the "higher" dicots. The Asteridae sens. lat. include the Ericales, Cornales, and Apiales in addition to the Asteridae sens. str. Two major lineages within the Asteridae sens. lat. are identified: the Dipsacales, Apiales, Asterales, and Campanulales in one, and the Gentianales, Scrophulariales, Lamiales, Boraginales, and Solanales in the other. This analysis demonstrates the utility of molecular phylogenies to help place problematic taxa, such as the Menyanthaceae, Oleaceae, and Callitrichaceae, within the Asteridae. Implications from this phylogenetic analysis and evidence from the fossil record lead to the suggestion that the origin and diversification of the major higher-dicot lineages occurred during a relatively short period of time about 80–95 million years ago.

The modern concept of the Asteridae, sensu Takhtajan (1980) and Cronquist (1981), is derived from the ancestral Monopetalae (de Jussieu, 1789) and Gamopetalae (de Candolle, 1813) by the elimination of many groups of plants bearing the original defining feature of fused corollas (Wagenitz, 1992). Cronquist (1981: 852) stated that "the Asteridae are the most advanced subclass of dicotyledons." This statement puts into words a generally held perception, based on traditional assumptions regarding trends in character evolution in the angiosperms, that the subclass is of relatively recent origin compared to other major groups of dicots (Sporne, 1969, 1975; Stebbins, 1974).

There is no consensus of opinion concerning the monophyly of the Asteridae. Whereas a combination of floral and embryological characters seems to define a natural group, portrayed as monophyletic in the treatments of Cronquist (1981) and Takhtajan (1980, but not 1987), chemical characters suggest two separate asterid lineages, each derived independently from ancestors in the Rosi-dae (Dahlgren, 1980). As Wagenitz (1977) pointed out, no division of the Asteridae into separate lineages can be constructed without having to postulate parallel evolution in morphology, embryology, and phytochemistry. Parsimony-based methods of phylogeny reconstruction offer a means of assessing phylogenetic information in which parallelisms exist, by establishing objective criteria for accepting one hypothesis of relationships (i.e., tree) over another hypothesis. Parsimony-based phylogeny reconstructions among major groups in the dicots are few. Donoghue & Doyle (1989), in their analysis of basal angiosperm lineages, identified a "higher"-dicot clade (i.e., derived relative to the basal dicots). This clade, to which all Asteridae, Rosidae, Dilleniidae, Caryophyllidae, and Hamamelidae, as well as certain members of the Magnoliidae, belong is characterized by the presence of tricolpate pollen. Hamby & Zimmer (1991) conducted a parsimony analysis of nuclear ribosomal RNA sequences in angiosperms and other seed plant groups, but found little resolution among

TABLE 1. Sources of *rbc*L sequences. Arrangement follows Cronquist (1981).

Family	Species	Source/voucher[a]
Monocots		
Liliidae		
Burmanniaceae	*Burmannia biflora*	MWC (unpublished)
Liliaceae	*Lilium superbum*	MWC (unpublished)
Orchidaceae	*Oncidium excavatum*	MWC (unpublished)
Commelinidae		
Poaceae	*Cenchrus setigerus*	Doebley et al. (1990)
	Puccinellia distans	Doebley et al. (1990)
Dicots		
Magnoliidae		
Lauraceae	*Persea americana*	Golenberg et al. (1990)
Magnoliaceae	*M. macrophylla*	Golenberg et al. (1990)
Nelumbonaceae	*Nelumbo lutea*	Les et al. (1991)
Nymphaeaceae	*Nuphar variegata*	Les et al. (1991)
	Nymphaea odorata	Les et al. (1991)
Hamamelidae		
Cercidiphyllaceae	*Cercidiphyllum japonica*[b]	*RGO s.n.*
Platanaceae	*Platanus racemosa*	EMG (unpublished)
Caryophyllidae		
Amaranthaceae	*Amaranthus hypochondriacus*	Michalowski et al. (1990)
Caryophyllaceae	*Stellaria media*	JHR, JRM & HDW (unpublished)
Chenopodiaceae	*Spinacia oleracea*	Zurawski et al. (1981)
Phytolaccaceae	*Phytolacca americana*	JHR, JRM & HDW (unpublished)
Plumbaginaceae	*Plumbago capensis*	DEG et al. (unpublished)
Polygonaceae	*Rheum* ×*cultorum*	DEG et al. (unpublished)
Dilleniidae		
Brassicaceae	*Brassica campestris*	JMN (unpublished)
Ericaceae	*Rhododendron hippophaeoides*	MWC & KK (unpublished)
Fouquieriaceae	*Fouquieria splendens*[b]	*Matthaei BG 860162*
Malvaceae	*Gossypium hirtum*	Gulov et al. (1990)
Violaceae	*Viola soraria*[b]	RGO (no voucher)
Rosidae		
Apiaceae	*Coriandrum sativum*[b]	JDP (no voucher)
Araliaceae	*Hedera helix*[b]	*RKJ s.n.*
Cornaceae	*Cornus mas*	Donoghue et al. (1992)
Fabaceae	*Medicago sativa*	Aldrich et al. (1986)
Grossulariaceae	*Brexia madagascarensis*	Soltis et al. (1990)
Hydrangeaceae	*Carpenteria californica*	Soltis et al. (1990)
Linaceae	*Linum perenne*	MWC (unpublished)
Onagraceae	*Clarkia xantiana*	KJS & EC (unpublished)
Polygalaceae	*Securidaca diversifolia*	MWC (unpublished)
Saxifragaceae	*Heuchera micrantha*	Soltis et al. (1990)
	Parnassia fimbriata	Soltis et al. (1990)
	Penthorum sedoides	Soltis et al. (1990)
Vochysiaceae	*Qualea* sp.	MWC (unpublished)
Asteridae		
Apocynaceae	*Apocynum cannabinum*[b]	RGO (no voucher)
Asteraceae	*Barnadesia caryophylla*	Michaels et al. (in prep.)
Bignoniaceae	*Catalpa* sp.[b]	*CWD s.n.*
Boraginaceae	*Borago officinalis*[b]	RGO (no voucher)
Callitrichaceae	*Callitriche heterophylla*[b]	*TCP 2152*
Calyceraceae	*Boopis anthemoides*	Michaels et al. (in prep.)

Volume 79, Number 2
1992

Olmstead et al.   251
Monophyly of the Asteridae Inferred
from *rbc*L Sequences

TABLE 1.   Continued.

Family	Species	Source/voucher[a]
Campanulaceae	*Campanula ramosa*	Michaels et al. (in prep.)
Caprifoliaceae	*Symphoricarpos albus*[b]	*RGO s.n.*
	Viburnum acerifolia	Michaels et al. (in prep.)
Convolvulaceae	*Convolvulus tricolor*[b]	RGO (no voucher)
Dipsacaceae	*Dipsacus sativus*	Michaels et al. (in prep.)
Gentianaceae	*Exacum affine*[b]	*Matthaei BG s.n.*
Hydrophyllaceae	*Hydrophyllum virginiana*[b]	RGO (no voucher)
Lamiaceae	*Lamium purpureum*[b]	RGO (no voucher)
Menyanthaceae	*Villarsia calthifolia*[b]	*RO 9726*
Oleaceae	*Ligustrum vulgare*[b]	*RGO s.n.*
Polemoniaceae	*Polemonium reptans*[b]	*BBG s.n.*
Scrophulariaceae	*Antirrhinum majus*[b]	*CWD s.n.*
Solanaceae	*Nicotiana tabacum*	Lin et al. (1986)
Valerianaceae	*Valeriana officinalis*	Michaels et al. (in prep.)
Verbenaceae	*Clerodendrum fragrans*[b]	*Matthaei BG 840210*

[a] BBG = Beal Botanical Garden, Michigan State University, CWD = Claude dePamphilis, DEG = David Giannasi, EC = Elena Conti, EMG Ed Golenberg, HDW = High Wilson, JDP = Jeffrey Palmer, JMN = Jackie Nugent, JHR = Jeff Rettig, JRM = James Manhart, KJS = Ken Sytsma, KK = Kathy Kron, Matthaei BG = Matthaei Botanic Garden, University of Michigan, MWC = Mark Chase, RKJ = Robert Jansen, RGO = Richard Olmstead, RO = Robert Ornduff, TCP = Thomas Philbrick.
[b] Sequences determined for this study.

the higher dicots. The higher dicots (sensu Donoghue & Doyle, 1989) are currently the subject of a morphology-based parsimony analysis aimed at identifying the relatives of the Asteridae (Hufford, 1992), and the Asteridae are the subject of a parsimony analysis of restriction sites in the cpDNA inverted repeat (Downie & Palmer, 1992).

The conceptual basis of our research into the phylogeny of the Asteridae is to develop a molecular data set derived from cpDNA sequences to address questions relating to the origin and diversification of the Asteridae. To do so requires sampling in sufficient depth among the entire higher dicots, as well as including representative outgroups from the "lower" dicots and monocots to root the resulting tree. Parsimony analysis of DNA sequence data is sensitive to taxonomic sampling. To prevent the attraction of distantly related branches on a parsimony tree, adequate sampling is necessary (see below). The choice of the chloroplast gene *rbc*L for our phylogenetic analysis of the Asteridae is based on prior studies (Palmer et al., 1988; Soltis et al., 1990; Golenberg et al., 1990; Doebley et al., 1990; Olmstead et al., 1990; Michaels et al., in prep.; Kim et al., 1992), which have revealed an appropriate amount of sequence variability at this phylogenetic level. Our analysis also benefits from the fact that *rbc*L is presently being sequenced in numerous groups of angiosperms, so that representative sequences outside the Asteridae are available as outgroups.

We undertook the study of Asteridae phylogeny with several goals in mind: (1) to test the monophyly of the Asteridae; (2) to identify major lineages of the Asteridae; (3) to evaluate ordinal circumscriptions; (4) to determine relationships among orders and among families within orders; (5) to help place taxa that are placed ambiguously in existing classifications; (6) to provide a basis for interpreting character evolution within the Asteridae; and (7) to provide a basis, along with a reassessment of traditional taxonomic characters, for a revised classification of the Asteridae. The present analysis will focus on the first two of these goals and will demonstrate the potential of molecular phylogenies for resolving ambiguously placed taxa. A second analysis, currently underway with greater sampling in the Asteridae, emphasizes familial and ordinal relationships. Suggestions for taxonomic revisions will await its outcome. The value of this *rbc*L-based phylogeny of the Asteridae for identifying family-level sister groups to aid in phylogenetic studies of specific families or orders is demonstrated elsewhere in this volume (Donoghue et al., 1992; Olmstead & Palmer, 1992).

MATERIALS AND METHODS

Plant material was either field-collected or obtained as fresh leaf material or seed from various sources (Table 1). DNA was isolated from fresh leaf material as either total cellular DNA following the modified CTAB procedure (Doyle & Doyle, 1987), or as purified chloroplast DNA (cpDNA)
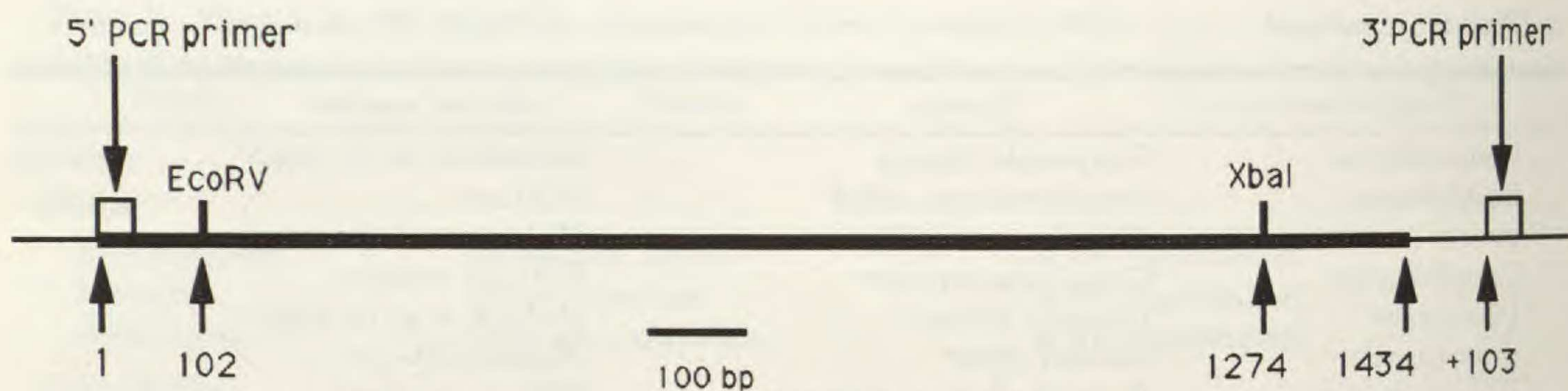
FIGURE 1. Cloning and PCR amplification strategy for *rbc*L. Length of the coding sequence (1434 bp) and position of the 3' PCR primer (beginning +103 bp from termination of coding region), shown here as in *Nicotiana tabacum*, are variable among species examined.

by the sucrose gradient method (Palmer, 1986). All cpDNAs and most total DNAs were further purified by CsCl/ethidium bromide gradient centrifugation.

An approximately 1550 bp segment of double-stranded DNA containing the complete coding sequence for the chloroplast gene *rbc*L was amplified using the Taq polymerase-mediated polymerase chain reaction (PCR). Two synthetic oligonucleotides were designed for use as amplification primers. The 5' primer is based on the first 26 nucleotide positions of the *rbc*L coding sequence and is two-fold degenerate at position 18 to account for the only difference between the maize and tobacco sequences in this region. The 3' primer is based on a 24 bp sequence that contains part of a stem-loop structure beginning 103 bp beyond the coding sequence termination for *rbc*L in tobacco (Fig. 1).

The sequence for *rbc*L was obtained following cloning of the PCR product for most of the species examined. However, the sequence for *Cercidiphyllum japonicum* and portions of several other sequences were obtained by direct sequencing of the PCR product following amplification with one biotin-labeled primer and strand separation on a streptavidin-agarose column (Mitchell & Merril, 1989). A cloning strategy was adopted that made use of highly conserved EcoRV and Xbal restriction sites (recognition sites at nucleotide positions 103–108 and 1269–1274, respectively) and the residual activity of the thermostable Taq polymerase in the PCR mixture (Fig. 1). Crude PCR product was digested with the two enzymes for two hours at 37°C and ligated into the plasmid cloning vector BlueScript Sk+ (Stratagene, Inc.), which had been digested previously with EcoRV to enable the cloning of blunt-end double-stranded DNA fragments. The residual Taq polymerase activity and nucleotide pool in the crude PCR product allow the filling-in of the four-base, 5' overhang at the Xbal restriction site at the same time that the restriction digest is occurring. Following transformation, clones

were selected with inserts of approximately 1200 bp (EcoRV-Xbal fragments) and 300 bp (Xbal-3' PCR primer fragments). Sequencing was accomplished using the dideoxy method with primers provided by G. Zurawski. For most taxa at least two independent clones of each fragment were sequenced to minimize possible errors introduced by misincorporation during PCR. Estimated rates of misincorporation in independent clones from the same PCR reaction ranged from 0 to 0.3%. This figure is in accord with published rates of PCR misincorporation (Saiki et al., 1988). Positions at which misincorporation was detected for a specific taxon were entered as unknown in the sequence for that taxon in the phylogenetic analysis.

A total of 1305 bp of sequence was compared beginning at position 103 and ending at position 1407. In the first 1407 bp of sequence, no insertions or deletions were observed in any of the sequences studied. Alignment of sequences was done by comparison to the sequence for *Nicotiana tabacum*. Beyond position 1407 sequence divergence becomes great, and small insertions and deletions make the alignment of homologous positions uncertain. Additional sequences were obtained from published and unpublished sources (Table 1). All sequences are complete for the 1305 bp region being compared with the exception of *Nelumbo*, *Nuphar*, and *Nymphaea* (1053 bp each) and *Platanus* (1163 bp). Of the 1305 bp compared, 612 positions are variable and 415 of these are phylogenetically informative. Parsimony analyses were performed using PAUP version 3.0n (Swofford, 1989) on a Macintosh IIfx computer using the heuristic search option with global branch swapping, MULPARS, and ten replicate runs with random order of taxon entry to search for the shortest trees. A bootstrap analysis (Felsenstein, 1985) of 100 replicates was performed using global branch swapping, MULPARS, and the CLOSEST addition sequence to assess the relative support of clades identified by the parsimony analysis. To facilitate

Volume 79, Number 2
1992

Olmstead et al.   253
Monophyly of the Asteridae Inferred
from *rbc*L Sequences

the bootstrap analysis of such a large data set, the topological constraint option in PAUP was used to constrain certain taxonomic groupings that had been identified as monophyletic by preliminary analyses. This approach effectively reduces the number of terminal taxa in the analysis, while maintaining all of the sequences, thereby enabling the optimal assessment of character state transformations over the tree. Internal nodes on the tree, where branching pattern is critical to the questions addressed by this analysis, were left unconstrained. Portions of the tree that were constrained in the bootstrap analysis are indicated in Figure 3.

RESULTS

Sequences were obtained for 57 taxa (Table 1), including 15 published sequences, 23 unpublished sequences (provided by M. Chase, D. Les, K. Sytsma, E. Golenberg, H. Michaels, J. Nugent, J. Manhart & D. Gianassi), and 19 sequences generated for this study. All sequences generated as part of this study are deposited with Genbank and are available upon request from the authors (direct requests to R. Olmstead). Sampling focused on the Asteridae, with additional sequences obtained for taxa selected because they are putatively closely related to the Asteridae and because they fill gaps among other dicot lineages. Outgroup sampling reflects to a great degree the diversity of groups presently being examined for *rbc*L sequences.

A Wagner parsimony analysis, in which all inferred nucleotide substitutions are equally weighted, yielded 16 minimum length trees of 2,638 steps and a consistency index (CI) of 0.29 (Kluge & Farris, 1969), from which a strict consensus tree was constructed (Fig. 2). The monocots and "lower" dicots (i.e., Magnoliideae with nontricolpate pollen) provide a good selection of outgroup taxa with which to root the portion of the tree representing the higher dicots, even though no outgroup for the angiosperms as a whole was included in the analysis.

The strict consensus tree (Fig. 2) is rooted arbitrarily using the woody Magnoliidae taxa, *Magnolia* and *Persea*, as outgroups. The tree shows three clusters of taxa among the lower dicots and monocots: (1) woody Magnoliidae (if rooted elsewhere among the lower angiosperms, these taxa form a clade), (2) monocots, and (3) Nymphaeales (minus *Nelumbo*). Two of these groups have been proposed recently to represent the basal branches of angiosperm phylogeny (woody Magnoliidae—Donoghue & Doyle, 1989; Nymphaeales—Hamby & Zimmer, 1991). Rooting the *rbc*L tree with the

woody magnoliids (Fig. 2) agrees with traditional angiosperm classification (Cronquist, 1981) and has been suggested by Donoghue & Doyle (1989) on the basis of a phylogenetic analysis using conventional taxonomic characters. Rooting the *rbc*L tree with the Nymphaeales (*Nuphar* and *Nymphaea*), as suggested by Hamby & Zimmer (1991) on the basis of rRNA sequence data, yields a tree (not shown) in which monocots and the woody Magnoliidae form a clade and the remaining dicots form another. Using either rooting, the higher dicots form a group corresponding to the "tricolpate" clade of Donoghue & Doyle (1989).

Five major clades are identified among the higher dicots (Figs. 2, 3). The basal branch consists of *Nelumbo* (Magnoliidae) and *Platanus* (Hamamelidae). The separation of *Nelumbo* from the rest of the water lilies (e.g., Nymphaeales) has been suggested by Donoghue & Doyle (1989) based on morphology, and by Les et al. (1991) based on *rbc*L sequences. The remaining taxa fall into four recognizable groups, but with only weak bootstrap support for any specific branching order among them (Figs. 3, 4). One group consists of the Saxifragaceae sens. str., represented by *Heuchera* and *Penthorum*. The majority rule consensus tree and bootstrap analysis (Fig. 3) suggest that *Cercidiphyllum* falls within or near this group. The most strongly supported group among the higher dicots is the Caryophyllidae, which occurred in 97% of the bootstrap replicates (Fig. 3). The remaining taxa form two main clades, one comprising most of the representatives of the Rosidae and Dilleniidae and the other predominantly of Asteridae. A bootstrap analysis (Fig. 3) provides relative estimates of support for the groupings in the critical region of the tree where the higher-dicot clades diverge.

Within the asterid clade, two main lineages are shown. One includes the orders Gentianales, Solanales, Boraginales, Scrophulariales, and Lamiales, while the other includes the Asterales, Campanulales, Dipsacales, Goodeniales, Apiales, and Menyanthaceae, all Asteridae sensu Cronquist (1981) except the Apiales (Rosidae). Two smaller clades are associated with the Asteridae near the base of the two main lineages. One of these clades consists of *Cornus* and *Carpenteria*, and the other includes *Fouquieria*, *Polemonium*, and *Rhododendron*.

In addition to the analysis of 57 *rbc*L sequences (Figs. 2–4), a larger preliminary analysis of 92 sequences, some of which were incomplete at the time of the analysis, was conducted (results not shown). This global analysis differed only slightly
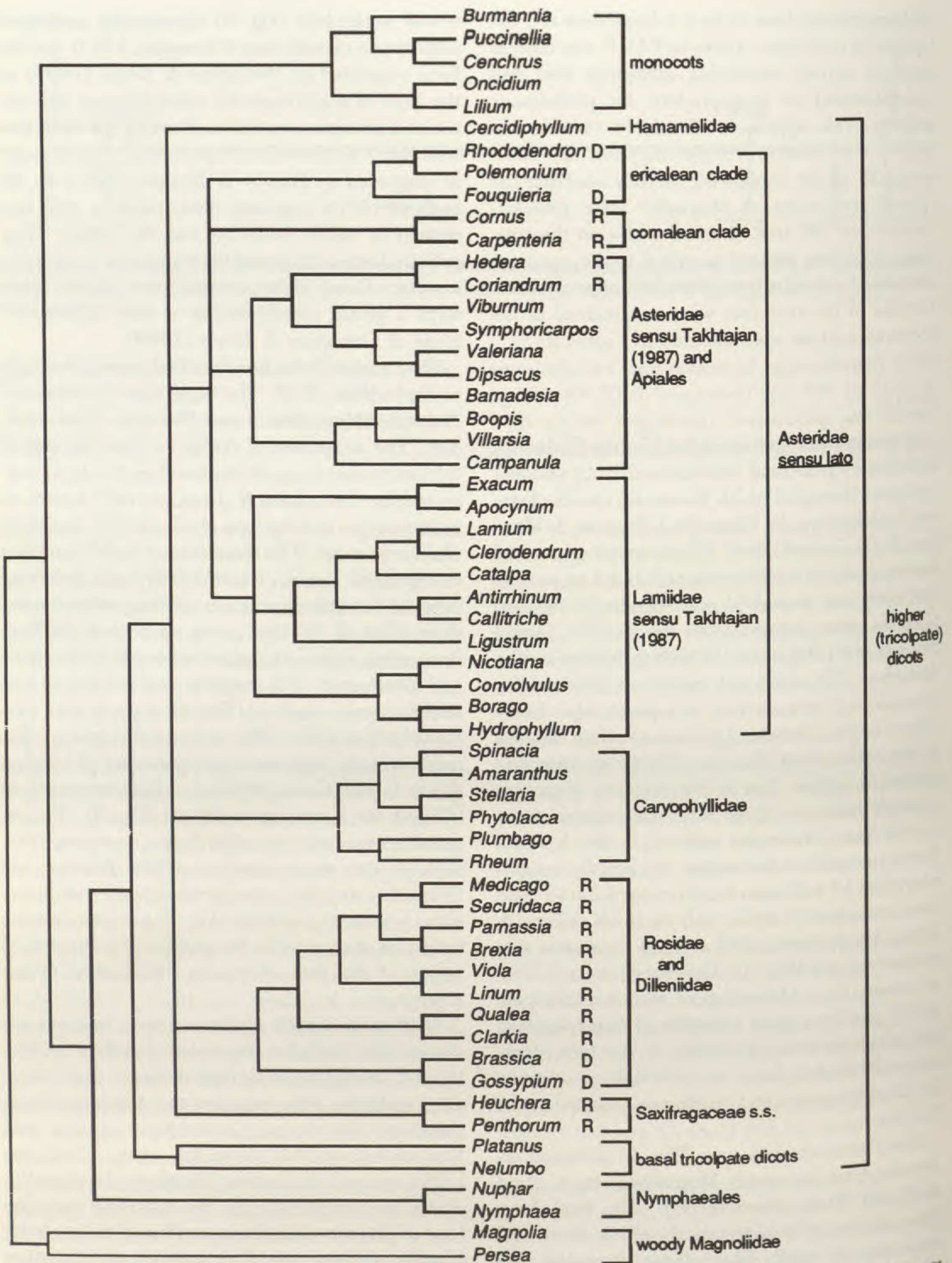
FIGURE 2. Strict consensus of 16 minimum length Wagner trees based on *rbc*L sequences (length = 2,638, CI = 0.29). The tree is arbitrarily rooted using the woody Magnoliidae as the outgroup. Representatives of the Rosidae (R) and Dilleniidae (D) are identified.

Volume 79, Number 2
1992

Olmstead et al.
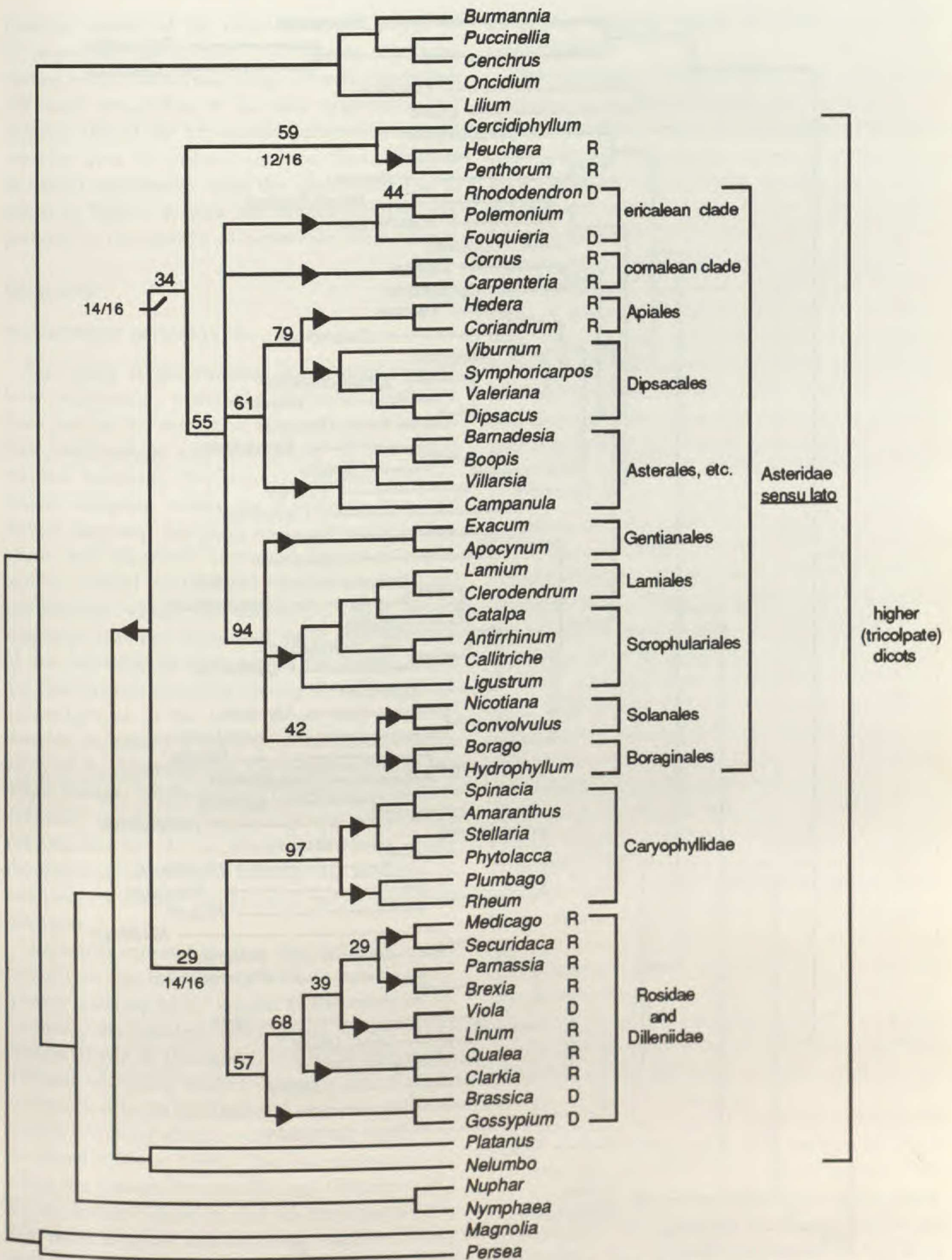Monophyly of the Asteridae Inferred
from *rbc*L Sequences

255

FIGURE 3. Results of a bootstrap analysis. Branching order in clades distal to arrows, in the direction arrows are pointing, is constrained topologically (except the ericalean clade, in which order was unconstrained). Major groups in the Asteridae sens. lat. are identified. Numbers above internodes of unconstrained portions of the tree indicate the percentage of bootstrap replicates supporting the distal clade. Fractions below internodes indicate the number of the 16 equally most parsimonious trees exhibiting that branch; all other clades are found in all 16 trees. Representatives of the Rosidae (R) and Dilleniidae (D) are identified.
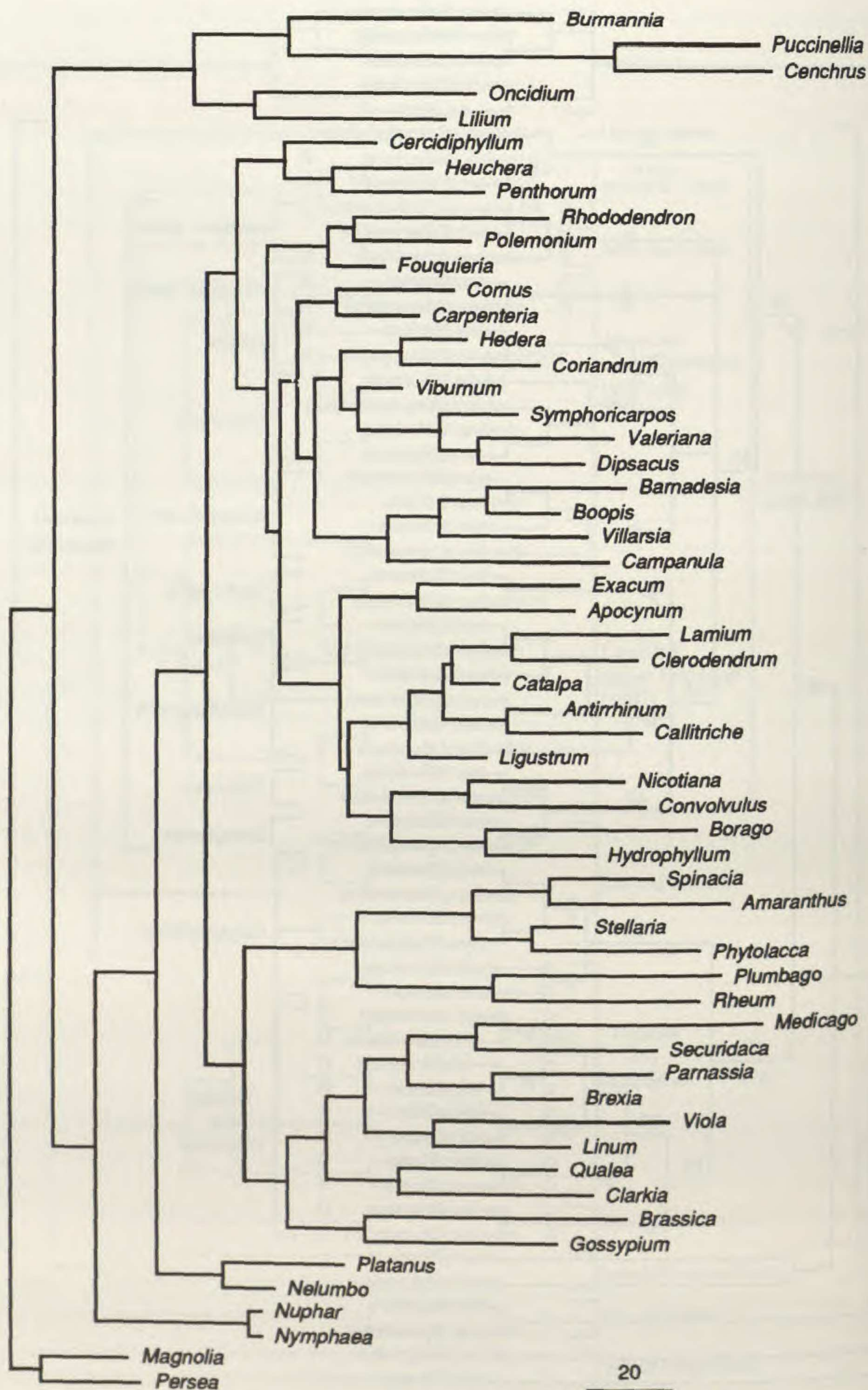
FIGURE 4. One of the 16 equally parsimonious Wagner trees based on *rbc*L sequences (length = 2,638, CI = 0.29). This tree is the one most similar to the preliminary analysis of 92 *rbc*L sequences (see text). Branch lengths are proportional to the number of inferred nucleotide substitutions (note scale at bottom).

Volume 79, Number 2
1992

Olmstead et al.   257
Monophyly of the Asteridae Inferred
from *rbc*L Sequences

from the results of the more detailed analysis of
57 sequences shown here. A complete report in-
cluding more detailed sampling within the Asteridae
will await completion of the data acquisition and
analysis. Out of the 16 equally parsimonious trees
resulting from the present analysis, the one that is
in closest agreement with the global analysis is
shown in Figure 4, with the branch lengths pro-
portional to the number of nucleotide substitutions.

DISCUSSION

PHYLOGENETIC IMPLICATIONS

Our goals of determining family and ordinal-
level relationships within the Asteridae depend in
large part on the resolution of questions of higher-
level relationships and monophyly of the entire
subclass Asteridae. Preliminary analyses with ex-
tensive sampling within the Asteridae, but with
limited outgroup sampling (e.g., *Spinacia*, *Heu-
chera*, and *Magnolia* alternatively and in combi-
nation), yielded contradictory hypotheses of basal
relationships within the Asteridae. These results
suggested that the origin and early diversification
of the Asteridae lie deep within the higher dicots
and that greater sampling among the higher dicots
and outgroups in the lower dicots and monocots
would be necessary to resolve the question of mono-
phyly of the Asteridae and the relationships among
major lineages of Asteridae. Therefore, it became
necessary to address the question of monophyly of
the Asteridae first, by sampling more broadly among
the angiosperms, before turning our attention to
relationships among families and orders within the
Asteridae.

An initial attempt to root the resulting angio-
sperm tree was made using *Welsitschia* (*rbc*L se-
quence courtesy of G. Furnier), a member of the
Gnetales, the putative sister group to the angio-
sperms (Doyle & Donoghue, 1986). However, the
inclusion of such a remote outgroup introduced a
systematic error in the form of a very long outgroup
branch attaching along a very long ingroup branch
(Swofford & Olsen, 1990). This long ingroup branch
led to the grasses *Puccinellia* and *Cenchrus* (Fig.
4), an unlikely place to root an angiosperm tree.
This result suggests that the much longer internode
distances among seed plant groups, relative to those
within the angiosperms, may require using some
form of a priori weighting of *rbc*L sequences among
seed plant lineages (e.g., weighting first and second
codon positions more than third positions or weight-
ing transversions over transitions). In contrast,
within the more closely related and more densely
sampled angiosperms, an equally weighted Wagner

parsimony analysis should be sufficient to infer
relationships.

Our approach was to restrict our sampling to
angiosperms on the assumption that adequate tax-
onomic sampling within the angiosperms would en-
able a Wagner parsimony analysis to produce a
better estimate of branching topology than would
be produced when one or more very long outgroup
branches were added. By this approach, the pattern
of branching among the higher dicots should be
most accurately estimated, regardless of the po-
sition of the root among the lower dicots or mono-
cots. The resulting tree can still provide insight
into the basal relationships within the angiosperms
by examining trees rooted on the basis of hypoth-
eses generated by other lines of evidence. If the
tree is rooted by the branch connecting the woody
magnoliids to the rest of the tree (Donoghue &
Doyle, 1989), then the monocots and the remain-
ing dicots each form monophyletic groups (Fig. 2).
If the tree is rooted by the branch connecting the
Nymphaeales to the rest of the tree (Hamby &
Zimmer, 1991), then the woody magnoliids and
the monocots form a clade and the remainder of
the dicots forms a clade (tree not shown). Neither
of these rootings imply relationships among the
basal angiosperms that are in complete agreement
with the other hypotheses of relationship cited above,
perhaps because sampling among the basal angio-
sperm lineages is incomplete. However, the higher-
dicot clade characterized by the presence of tri-
colpate pollen, which was identified by Donoghue
& Doyle (1989), is supported by the *rbc*L sequence
analysis. Because the focus of this study is on the
higher dicots, the Asteridae in particular, sampling
among the basal lineages is sporadic and far from
representative of the important basal angiosperm
groups. However, sampling among the basal an-
giosperms should be sufficient to provide outgroup
comparison to assess relationships among the high-
er dicots where sampling is more complete, es-
pecially within the Asteridae.

Relationships among the higher (i.e., tricolpate)
dicots implied by the *rbc*L tree include an initial
branch consisting of *Nelumbo* and *Platanus*. This
is an unexpected pairing, but many other poten-
tially closely related taxa in the Hamamelidae and
Magnoliidae are not included and both of these
sequences are incomplete (*Nelumbo* lacks 252 bp;
Platanus lacks 142 bp). Nevertheless, their place-
ment near the base of the tricolpate dicots is con-
sistent with other evidence (Donoghue & Doyle,
1989).

Perhaps the most striking implication of the *rbc*L
tree is that the major higher-dicot lineages, in-

cluding the Caryophyllidae, Rosidae/Dilleniidae, and the "advanced" subclass Asteridae, trace to an apparent radiation point early in the evolution of the higher dicots. The lack of resolution among the major higher-dicot lineages is due to the small number of nucleotide substitutions between branch points relative to the total length of time since the lineages diverged. The observation of few substitutions between branch points, combined with an assumption of relative constancy of substitution rate, implies a rapid radiation of major higher-dicot groups. The lack of resolution also stems, in part, from homoplasy at nucleotide positions that are informative regarding relationships at that level. There is hope that sufficient information does exist in the rbcL sequences to offer greater resolution than this study provides, by sampling additional taxa near the critical juncture of the higher-dicot radiation (see discussion on sampling below).

Three major higher-dicot lineages are identified which coincide, more or less, with the taxonomic level of subclass in the classifications of Cronquist (1981) and Takhtajan (1980). In addition, a fourth smaller group is identified comprising several members of the Saxifragaceae sens. str. (Rosidae) and Cercidiphyllum (Hamamelidae). The Saxifragaceae sens. lat. were the subject of a recent rbcL sequence analysis (Soltis et al., 1990), in which it was concluded that the traditional, broadly defined Saxifragaceae are a heterogeneous group of more or less distantly related taxa. However, the restricted sampling of that study (see discussion below) limited its ability to interpret which taxa belong together in a more narrowly defined Saxifragaceae. By including representatives of the Soltis et al. (1990) study in the larger analysis reported here, Saxifragaceae sens. str. represented by Heuchera and Penthorum (Fig. 2), along with Astilbe and Itea (results not shown), are identified. The association of Cercidiphyllum with this group of Saxifragaceae, either as separate lineages connected at the same unresolved node on the main higher-dicot lineage (Fig. 2) or together as a clade (Figs. 3, 4), suggests an association between the woody Hamamelidae (e.g., Cercidiphyllum) and basal Rosidae (e.g., Saxifragaceae). This association may be a key to understanding higher-dicot diversification, but much greater sampling of taxa in these groups is needed before a clear picture can emerge.

The Caryophyllidae, including Polygonaceae and Plumbaginaceae, appear as a strongly supported, monophyletic group, whose circumscription in traditional classifications coincides completely with the molecular evidence presented here and in the more detailed analysis of the Caryophyllales of Rettig et

al. (in prep.). Our results indicate that the Polygonaceae and Plumbaginaceae form a clade that is the sister group to the Caryophyllales as has been suggested by Rodman et al. (1984). The Caryophyllidae are consistently supported by numerous synapomorphies on all the equally parsimonious trees and are supported by the highest bootstrap value (97%) of any of the major higher-dicot lineages identified by this analysis (Fig. 3).

The Dilleniidae and Rosidae form a second major lineage of higher dicots (with several important exclusions). The results of this analysis suggest that neither subclass, sensu Takhtajan (1980) and Cronquist (1981), is a monophyletic group (Figs. 2, 3). In addition, representatives of each of these subclasses occur in the third major higher-dicot lineage, which consists primarily of the Asteridae, and several members of the Saxifragaceae sens. lat. (Rosidae) form a distinct lineage at or near the base of the higher-dicot diversification. The ten taxa that form the Rosidae/Dilleniidae clade in our analysis are a dramatic underrepresentation of the diversity in these two subclasses. However, the existence of this clade, as well as the lack of any clear distinction between the representatives of these two subclasses in the rbcL tree, is supported by preliminary evidence from other workers sequencing rbcL in the Rosidae and Dilleniidae (M. Chase, pers. comm.; R. Price, pers. comm.).

The third major clade of higher dicots is the Asteridae sensu lato, which include several taxa traditionally placed in the Rosidae or Dilleniidae. Two minor clades at the base of the Asteridae (Cornus and Carpenteria in one, and Rhododendron, Fouquieria, and Polemonium in the other) are unresolved with respect to the divergence of the two main lineages of Asteridae. The presence of these two clades reflects a grade in the evolution of the Asteridae recognized by many previous treatments (although evidence from the preliminary analysis with more taxa suggests that one of these small groups, Cornus and Carpenteria, may belong with the Dipsacales and Asterales). Cronquist (1981) and Takhtajan (1980) recognized a narrowly defined Cornales, placed in the Rosidae, but suggested that the Asteridae may have arisen from a cornalean ancestor or share a close common ancestor with the Cornales. Thorne (1983) and Dahlgren (1980) recognized the more inclusive superorder, Corniflorae, comprising the Cornales and Dipsacales, and including either the Apiales (Thorne, 1983) or the Ericales and Fouquieriaceae (Dahlgren, 1980). In both of their treatments this group is placed in a position of lesser "advancement" relative to the rest of the orders of the Asteridae.

Volume 79, Number 2
1992

Olmstead et al.   259
Monophyly of the Asteridae Inferred
from *rbcL* Sequences

None of the above-mentioned recent treatments of angiosperm classification place the Polemoniaceae with the Ericaceae and Fouquieriaceae, although Thorne (1983) placed the Fouquieriaceae near the Polemoniaceae and each of the other treatments cited above have placed the Fouquieriaceae near the Ericaceae. It is clear from the *rbcL* tree that the Cornales, Ericales, and Fouquieriaceae arose at an early period in the diversification that gave rise to the Asteridae sensu Cronquist (1981), after the separation of the entire lineage from the one leading to most other Rosidae and Dilleniidae. Additional molecular support for the inclusion of these two lineages in the Asteridae sens. lat. comes from the analysis of restriction site mapping of the cpDNA inverted repeat (Downie & Palmer, 1992) and further taxonomic sampling of *rbcL* sequences in the Ericales and related taxa (K. Kron & M. Chase, pers. comm.).

Two primary lineages emerge from the unresolved basal portion of the Asteridae. One corresponds to Takhtajan's (1987) subclass Lamiidae, including the orders Gentianales, Lamiales, Scrophulariales, Solanales, and Boraginales. This lineage is one of the most strongly supported clades in the higher dicots with a bootstrap value of 94% (Fig. 3) and is identified in the study of Downie & Palmer (1992). Four of these orders, Gentianales, Lamiales, Solanales, and Boraginales, appear to be monophyletic groups based on the limited sampling presented here. This tentative conclusion is supported by a preliminary analysis of a larger number of taxa within this clade (results not shown). The Lamiales form a monophyletic group with the Scrophulariales as suggested by Wagenitz (1992) and are not close to the Boraginaceae, in agreement with Cantino (1982). The Oleaceae, represented in this study by *Ligustrum*, have been placed alternately with the Gentianales (Dahlgren, 1980; Takhtajan, 1987) and the Scrophulariales (Cronquist, 1981) and are identified as one of several families of questionable placement in the Asteridae by Wagenitz (1992). This study suggests that the Oleaceae represent a basal branch of the clade leading to the Scrophulariales and Lamiales.

Although groups corresponding to orders can be identified within this primary lineage, relationships among these groups remain unclear. The association of the Lamiales with the Scrophulariales appears to be well supported, but the suggested relationship between the Solanales and the Boraginales is weakly supported by the bootstrap analysis (Fig. 3). In the preliminary analysis of 92 sequences (not shown), including greater sampling within this lineage, the Boraginales come out with the Scroph-

ulariales/Lamiales rather than with the Solanales. More extensive taxonomic sampling and more phylogenetically informative data may be necessary to resolve ordinal relationships within this lineage.

The second primary lineage of Asteridae to emerge from the early diversification of the subclass corresponds closely to the Asteridae sensu stricto of Takhtajan (1987), but with the inclusion of the Apiales. The bootstrap value of 61% (Fig. 3) indicates moderate support for this clade and reflects some support in the data for inclusion of the Cornales. Two clades are recognized within this lineage, one comprising the Apiales and Dipsacales and the other comprising representatives of the Asterales, Campanulales, Goodeniales, and Menyanthaceae (*Villarsia*). The close molecular association between the Apiales and the Dipsacales was predicted only by the classification of Thorne (1983) among recent angiosperm classifications, although a similarity in secondary chemistry between the Apiales and Asterales also has been noted (Hegnauer, 1977). The placement of an order of Rosidae (i.e., Apiales) sensu Cronquist (1981) well within the Asteridae will surprise many observers. Nevertheless, a bootstrap value of 79% (Fig. 3) shows relatively strong support for a sister-group relationship between the Apiales and Dipsacales. This relationship is supported also by the work of Downie & Palmer (1992) and of Hamby & Zimmer (pers. comm.) using rRNA sequences. In the Dipsacales, the Caprifoliaceae sens. lat. (represented by *Viburnum* and *Symphoricarpos*) appear to form a paraphyletic group from which the Valerianaceae and Dipsacaceae are derived. Wagenitz (1992) identifies the Dipsacales and the Campanulales/Asterales (including Goodeniaceae) as "good candidates" for monophyletic groups, and the *rbcL* analysis provides additional support for his view. More detailed analyses of both of these monophyletic groups based on *rbcL* data are reported elsewhere (Campanulales/Asterales—Michaels et al., in prep.; Dipsacales—Donoghue et al., 1992).

The placement of the Menyanthaceae in the Campanulales/Asterales clade is an unanticipated result, but one that illustrates a strength of molecular approaches to phylogeny reconstruction. The Menyanthaceae have been placed alternately in the Gentianales (e.g., Takhtajan, 1987) and Solanales (e.g., Cronquist, 1981), but Bohm et al. (1986) could not find support for either placement on the basis of flavonoid data. Pollard & Amuti (1981) recognized a similarity between the Menyanthaceae and the Campanulales/Asterales on the basis of a primary reliance on inulin as a storage compound, but also included other more distantly

related families (e.g., Boraginaceae). Considering these conflicting hypotheses, it is not surprising that Wagenitz (1992) considers the placement of the Menyanthaceae to be "still controversial." The *rbc*L sequence data place the Menyanthaceae (represented here by *Villarsia*) squarely in the Campanulales/Asterales clade, a placement confirmed by sequencing a second member of the family, *Menyanthes* (unpublished data) and by the cpDNA restriction site analysis of Downie & Palmer (1992).

Another controversially placed taxon included in this analysis is the aquatic plant *Callitriche*. With its very reduced flowers and modifications associated with the aquatic habit, *Callitriche* has been assigned to a position, based on gynoecial and embryological characters, in or near the Lamiales in most recent treatments (Dahlgren, 1980; Thorne, 1983; Takhtajan, 1987; but see Cronquist, 1981). The analysis of *rbc*L sequence data suggests that *Callitriche* belongs in the Scrophulariales/Lamiales clade, but more closely related to the Scrophulariaceae than to the Lamiales.

The uncertain placement of some taxa in classifications based on conventional sources of data (e.g., morphology, anatomy, and secondary chemistry) is often the result of divergent evolution in these characters. This obscures relationships because derived characters shared between close relatives may no longer be apparent. In these circumstances cpDNA sequence data may have their greatest influence on classifications, because the stochastic nature of nucleotide substitutions in cpDNA is not expected to be coupled with differing rates of evolution of conventional characters. Therefore, taxa ambiguously placed in traditional classifications should resolve on an *rbc*L tree as confidently as any other taxa. This does not imply that cpDNA sequence divergence cannot be unpredictably variable and that this variability cannot introduce error or uncertainty into phylogenetic analysis of cpDNA sequence data (Doebley et al., 1990; Swofford & Olsen, 1990).

FLORAL EVOLUTION AND THE FOSSIL RECORD

The reconstruction of a framework phylogeny of the higher dicots, the identification of a monophyletic Asteridae sens. lat., and the delineation of primary lines of descent within the Asteridae enable one to begin to evaluate hypotheses of character evolution within the higher dicots and Asteridae. The results of our *rbc*L sequence analysis concur with the phylogenetic analysis based on conventional characters by Donoghue & Doyle (1989) in the identification of a clade of higher dicots char-

acterized by the presence of tricolpate pollen. Whereas our sampling is more representative of the higher dicots and that of Donoghue & Doyle is more representative of the lower dicots and monocots, it is heartening to observe the concordance in results at the point where the two studies converge.

Among the higher, tricolpate dicots, evolution of floral morphology has proceeded from ancestors with numerous parts spirally arranged to a reduced, fixed number of floral segments arranged in whorl. The evolution of whorled floral appendages has probably occurred more than once in dicots and certainly arose independently in dicots and monocots, but our analysis of *rbc*L sequence data suggests that a single origin of this floral arrangement may be sufficient to explain its presence in the major groups of higher dicots (e.g., Caryophyllidae, Dilleniidae, Rosidae, Asteridae). More extensive sampling of cpDNA sequences among the early tricolpate dicot lineages (e.g., Ranunculales, Hamamelidales) will be needed to determine whether a single origin is, in reality, sufficient to explain the distribution of this character among the higher dicots.

The lack of resolution of relationships among the major lineages of higher dicots suggests that the origin and divergence of these groups occurred close together in time and probably soon after the evolution of whorled floral appendages. Flowers with all parts in whorls are known by the Cenomanian age, approximately 95 million years ago (Friis & Crepet, 1987). By the middle Late Cretaceous (Santonian–Campanian), whorled flowers "usually with the perianth and androecium in whorls of five, and gynoecium of two or three carpels" were dominant (Friis & Crepet, 1987) and the first sympetalous flowers are known (Friis, 1985). The fossil record, notoriously incomplete when it comes to flowers, is entirely consistent with a nearly simultaneous (in geologic terms) origin of the major higher-dicot groups. Even though no fossils of clearly asterid affinity, sensu Cronquist or Takhtajan, are known until the Tertiary, one implication of our phylogenetic analysis is that the origin of the Asteridae sens. lat. was close in time to the appearance of other higher-dicot groups in the Late Cretaceous (Doyle & Donoghue, 1986; Hennig, 1966; Marshall, 1990). Fossil fruits assignable to the extant genus *Cornus* (E. M. Friis, unpublished data, cited in Eyde, 1988) and fossil flowers and fruits referable to the Ericales (Friis, 1985) are known from the Late Cretaceous. Both of these fossils belong to orders, Cornales and Ericales, respectively, identified by the *rbc*L analysis to belong to

Volume 79, Number 2
1992

Olmstead et al.   261
Monophyly of the Asteridae Inferred
from *rbcL* Sequences

the Asteridae sensu lato. Therefore, claims that fossils of asterid affinity do not appear until the Tertiary (Mueller, 1981; Cronquist, 1981) reflect our misconception of asterid affinity rather than a lack of fossil evidence.

The identification of a clade comprising the Asteridae sens. lat. that originated early in the diversification of the higher dicots and which is characterized (in large part) by the fusion of perianth parts suggests that the innovation of perianth fusion occurred soon after the evolution of whorled floral appendages. This implication of our analysis may seem strikingly at odds with the traditional concept of floral evolution, namely, many, spirally arranged parts *to* few, whorled parts *to* fused parts, with a phylogenetic diversification at each stage. However, there is no contradiction with this traditional concept of floral evolution; the only difference is that the transition from the second to the third stages occurred in rapid succession and that the phylogenetic radiation within each of these two floral plans occurred *after* that transition.

The interpretation of floral evolution within the Asteridae sens. lat. poses interesting hypotheses concerning the evolution of perianth fusion. If the two lineages that do not exhibit perianth fusion (*Cornus/Carpenteria* and *Hedera/Coriandrum*) represent the retention of the ancestral state of free floral parts, then the fusion of floral parts must have occurred at least three times during the evolution of the Asteridae sens. lat.: (1) Dipsacales, (2) Asterales/Campanulales, and (3) a clade comprising the ericalean group and the Lamiales/Scrophulariales/Gentianales/Boraginales/Solanales (evidence from phylogenetic studies in the Ericales suggests a separate origin of perianth fusion in that order (Kron & Chase, pers. comm.)). Alternatively, if a single origin of perianth fusion is postulated to have occurred in the Asteridae sens. lat., then either two reversals would be necessary to explain the distribution of free perianths, if the Cornales belong on the lineage leading to the Dipsacales, Apiales, Campanulales, and Asterales, or only one reversal, if the Cornales belong at the base of the entire clade. A simple parsimony argument, in which reversals and parallel evolution are equally likely, would favor a single origin of perianth fusion in the Asteridae sens. lat. By implication, the Apiales and possibly the Cornales would represent groups in which the existence of a free perianth is derived from an ancestrally fused perianth. It is noteworthy that the only putative cases of reversal in perianth fusion (e.g., Cornales and Apiales) are postulated to have occurred early in the diversification of the Asteridae, perhaps be-

fore subsequent floral evolution acted to constrain the development of the perianth to be obligately fused (Donoghue, 1989). Some recent evidence from studies of corolla development in the Asteridae and the Apiales (Erbar, 1988, 1991) is consistent with the hypothesis presented here that the polypetalous condition in the Apiales may represent a reversal from an ancestral sympetalous state. Erbar (1991) identified two developmental patterns leading to sympetaly in the Asteridae. In early sympetaly the corolla is initiated as a ring from which petal lobes later develop, whereas in late sympetaly the corolla is initiated as distinct petals, which fuse later in development. Corolla is of the late type in the families of the Lamiidae and the early type in the orders Asterales sens. lat. and Dipsacales. Corolla development in the Apiales is initiated as a ring, as in the Dipsacales and Asterales, but its development ceases when petal lobes are initiated, resulting in a corolla of apparently free petals (a correlation between ring formation and inferior ovary may present an alternate explanation, L. Hufford, pers. comm.). Comparable observations on corolla development in the Cornales is unavailable.

SAMPLING

Sampling is a critical issue that often is given insufficient attention in molecular phylogenetic studies. Sampling can affect both the resolution of a phylogenetic analysis and the effectiveness of statistical evaluation of the results (e.g., bootstrap analysis). Both the number of characters (e.g., nucleotides) and the number of taxa are factors that influence the resolution and reliability of an analysis. Of these two sampling considerations, we perceive the issue of taxonomic sampling to be the more critical. An insufficient number of characters often will result in a lack of resolution, rather than incorrect topology, whereas insufficient taxonomic sampling may lead to incorrect topology. This is particularly likely when unequal rates of nucleotide substitution exist within a clade or when the amount of substitution in the sequences being compared is high relative to the phylogenetic distance between sampled taxa (Felsenstein, 1978; Swofford & Olsen, 1990). In a parsimony analysis, an uneven sampling of taxa within a clade can have the same effect as unequal rates among lineages within the clade. Likewise, too few taxa sampled in a clade, even if evenly distributed, can have the effect of raising the effective rate of substitution for the sequence being compared (i.e., substitutions per internode length on a cladogram). Swofford & Ol-
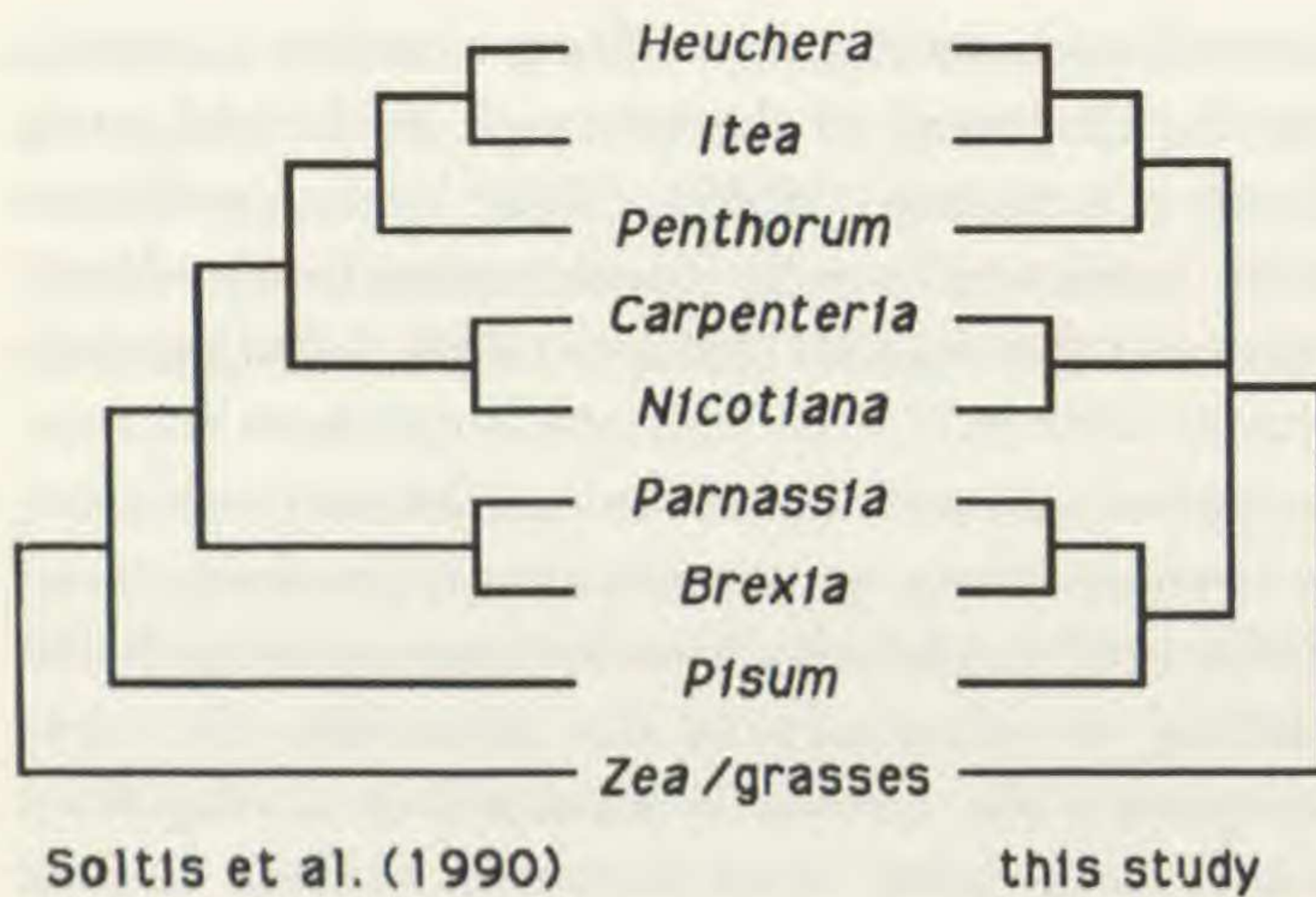
FIGURE 5. A comparison of the results of Soltis et al. (1990) and this study for taxa shared between the two studies.

sen (1990: 497) recommended that taxonomic sampling be carried out "so as to divide long branches reasonably evenly" and go on to point out that "long branches (sparse regions) within the ingroup can contribute to systematic errors," because in a parsimony analysis, "multiple substitutions are more easily detected in dense[ly sampled] regions."

Although increasing the taxonomic sampling density has the beneficial consequence of increasing the effectiveness of a parsimony analysis, it also has the paradoxical effect of reducing the apparent support for the resulting tree by the most commonly used means of evaluating tree branching patterns, the bootstrap (Felsenstein, 1985, and pers. comm.). The bootstrap approach provides a quantitative value for the occurrence of specific clades given the available data. Low bootsrap values can occur for two reasons: (1) when high levels of homoplasy exist in the data, making alternative groupings of taxa nearly equally likely; and (2) when there are few characters to support each branch point. Increasing the number of taxa without increasing the number of characters can increase the level of homoplasy in an analysis (Archie, 1989; Sanderson & Donoghue, 1989), and by splitting branches, will reduce the number of characters supporting each branch point, thereby reducing the bootstrap values on the resulting tree. Other methods of evaluating trees in a parsimony analysis, including decay analysis and comparing the distribution of all possible trees (Donoghue et al., 1992), are made more difficult, or rendered impossible, by increasing the number of taxa included in an analysis.

It is our concern that studies involving a small number of taxa drawn from a large and heterogeneous taxonomic class may show support for topologies that do not reflect phylogeny as accu-

rately as studies involving a greater sampling within the same taxonomic class. To illustrate this point, we offer two examples of studies involving small numbers of taxa drawn from the same taxonomic distribution as our analysis of 57 taxa reported here. First, anecdotal evidence from our analysis of *rbc*L sequences comes from the behavior of the sequence for *Spinacia*. In preliminary analyses with varying sets of taxa, but with no other members of the Caryophyllidae included, the sequence for *Spinacia* came out either within the Asteridae, within the Rosidae, or on a branch by itself at the base of the higher-dicot radiation. Only with the inclusion of other sequences from the Caryophyllidae (provided by D. Giannasi, J. Rettig & J. Manhart) did the position of *Spinacia* and the rest of the Caryophyllidae consistently come out near the base of the higher-dicot radiation (Fig. 2), regardless of which other taxa representing other clades were included in the analysis.

Soltis et al. (1990) employed *rbc*L sequence data to address the question of phylogenetic relationships within the Saxifragaceae sens. lat. and to determine if they represent an assemblage of relatively unrelated taxa. In addition to eight members of the Saxifragaceae sens. lat., three other diverse angiosperms, *Nicotiana* (Asteridae), *Pisum* (Rosidae), and *Zea* (monocot), were included in their analysis. Our analysis included five of their sequences from the Saxifragaceae sens. lat. as well as *Nicotiana*, *Medicago* (a close relative of *Pisum*), and two grass species in the same family as *Zea* (a sixth member of the Saxifragaceae sens. lat. and *Pisum* were also included in our preliminary analysis of 92 taxa). The topologies for the comparable portions of the tree from Soltis et al. (1990) and from our analysis differ primarily in the placement of the root (Fig. 5). In our view, the tree of Soltis et al. (1990) is rooted artifactually by the attraction of two long-branch lineages, *Pisum* and *Zea*. The resulting tree, otherwise topologically congruent with our tree, implies putatively erroneous conclusions concerning relationships within the Saxifragaceae sens. lat. and their position within the higher dicots. The source of the apparent error in their tree stems from: (1) the coincidence of including members of the Fabaceae (*Pisum*) and Poaceae (*Zea*), both fast-rate lineages for *rbc*L sequence evolution (note long branches in Fig. 4); and (2) rooting their analysis using the remote outgroup *Zea*, which has been shown to be diverging at a fast rate even when compared to other grasses (Doebley et al., 1990). This is not meant as a criticism of the study of Soltis et al. (1990), which was one of the first to use *rbc*L

Volume 79, Number 2
1992

Olmstead et al.   263
Monophyly of the Asteridae Inferred
from *rbc*L Sequences

sequences in a phylogenetic analysis, but rather as an example to illustrate the potential pitfall of inadequate taxonomic sampling.

SUMMARY AND PROSPECT

It is remarkable how many *rbc*L sequences have been acquired in the short time since the gene was first suggested for use in phylogenetic studies (Zurawski & Clegg, 1987; Ritland & Clegg, 1987; Palmer et al., 1988). Several recent studies (Golenberg et al., 1990; Soltis et al., 1990; Doebley et al., 1990; Michaels et al., in prep.; Les et al., 1991; Donoghue et al., 1992; Kim et al., 1992) and additional studies currently in progress (D. Giannasi, pers. comm.; J. Rettig & J. Manhart, pers. comm.; D. Clark, pers. comm.; M. Chase, pers. comm.) atest to the mushrooming interest in the application of *rbc*L sequences to phylogenetic studies in flowering plants. A recent survey of researchers (conducted in July 1990 by J. Palmer) revealed that over 200 *rbc*L genes had been sequenced; by the time this study is published, that number will have been greatly eclipsed. Much credit is owed to the generosity of G. Zurawski for providing sequencing primers to all interested researchers and to the open exchange of information on techniques and of sequences among researchers. The prospect in the near future for a comprehensive angiosperm phylogeny based on *rbc*L sequences is great.

The Asteridae sens. lat. have been the focus of the most extensive sampling for *rbc*L sequences, with more than 40 families and over 100 species sampled (Michaels et al., in prep.; Olmstead et al., this study, and unpublished; Kim et al., 1992; Donoghue et al., 1992; K. Kron & M. Chase, pers. comm.; D. Soltis, pers. comm.). The Asteridae provide an exemplary case study of how cpDNA sequence data can address systematic questions at a variety of levels, from the identification of a monophyletic, higher-order group, the Asteridae sensu lato (this study), to the circumscription of orders and relationship among families within orders (Michaels et al., in prep; Donoghue et al., 1992; Olmstead et al., unpublished), through to the study of tribal relationships within a large and diverse family, the Asteraceae (Kim et al., 1992).

The expanding body of evidence from *rbc*L studies does not mean that avenues for additional work are narrowing. To the contrary, *rbc*L sequence studies have identified many problems where the simple solution of increased taxonomic sampling will not suffice. The prospects for future research lie in several directions: (1) A slowly evolving gene

such as *rbc*L has a lower limit of effective resolution in phylogenetic studies (Doebley et al., 1990; Kim et al., 1992). Many problems within orders, families, or even large genera may be addressed more effectively by applying the techniques used in current *rbc*L studies to more rapidly evolving chloroplast genes, as well as by continuing the now widely used approach of cpDNA restriction site analysis. The former approach is currently underway in studies of the Asteraceae (R. Jansen, pers. comm.), Polemoniaceae (K. Steele, pers. comm.), and the Lamiales/Scrophulariales (R. Olmstead, unpublished). (2) Problems of ancient radiations (e.g., the higher-dicot radiation discussed above) may remain unresolved with *rbc*L sequences and will require additional data, perhaps 3–5 times as much cpDNA sequence to achieve resolution. (3) It should be noted that data from *rbc*L sequences are all derived from a single gene and may be subject to unknown evolutionary constraints. Phylogenetic reconstructions derived independently from chloroplast genes of differing functions (i.e., genes not involved in photosynthesis) or from mitochondrial or nuclear genes (e.g., rDNA genes) need to be conducted on the same taxa to confirm results and to test the underlying assumption that the nucleotide substitutions sampled from one gene represent a random sample. Overall, the prospect is bright that significant advances in our understanding of plant systematics will be forthcoming from continued studies of chloroplast and other DNA sequences.

LITERATURE CITED

ALDRICH, J., B. CHERNEY, E. MERLIN & J. D. PALMER. 1986. Sequence of the *rbc*L gene for the large subunit of ribulose bisphosphate carboxylase-oxygenase from alfalfa. Nucleic Acids Res. 14: 9534.

ARCHIE, J. W. 1989. Homoplasy excess ratios: new indices for measuring levels of homoplasy in phylogenetic systematics and a critique of the consistency index. Syst. Zool. 38: 253–269.

BOHM, B. A., K. W. NICHOLLS & R. ORNDUFF. 1986. Flavonoids of the Menyanthaceae: intra- and interfamilial relationships. Amer. J. Bot. 73: 204–213.

CANDOLLE, A. P. DE. 1813. Théorie Élémentaire de la Botanique. Paris.

CANTINO, P. D. 1982. Affinities of the Lamiales: a cladistic analysis. Syst. Bot. 7: 237–248.

CRONQUIST, A. 1981. An Integrated System of Classification of Flowering Plants. Columbia Univ. Press, New York.

DAHLGREN, R. 1980. A revised system of classification of the angiosperms. J. Linn. Soc., Bot. 80: 91–124.

DOEBLEY, J., M. DURBIN, E. M. GOLENBERG, M. T. CLEGG & D. P. MA. 1990. Evolutionary analysis of the large subunit of carboxylase (*rbc*L) nucleotide sequence among the grasses (Gramineae). Evolution 44: 1097–1108.

DONOGHUE, M. J. 1989. Phylogenies and the analysis of evolutionary sequences, with examples from seed plants. Evolution 43: 1137–1156.

———— & J. A. DOYLE. 1989. Phylogenetic studies of seed plants and angiosperms based on morphological characters. Pp. 181–193 in B. Fernholm, K. Bremer & H. Jornvall (editors), The Hierarchy of Life. Elsevier Science Publ., Amsterdam.

————, R. G. OLMSTEAD, J. F. SMITH & J. D. PALMER. 1992. Phylogenetic relationships of Dipsacales based on rbcL sequences. Ann. Missouri Bot. Gard. 79: 333–345.

DOWNIE, S. R. & J. D. PALMER. 1992. Restriction site mapping of the chloroplast DNA inverted repeat: a molecular phylogeny of the Asteridae. Ann. Missouri Bot. Gard. 79: 266–283.

DOYLE, J. A. & M. J. DONOGHUE. 1986. Seed plant phylogeny and the origin of angiosperms: an experimental cladistic approach. Bot. Rev. (Lancaster) 52: 321–431.

DOYLE, J. J. & J. L. DOYLE. 1987. A rapid isolation procedure for small quantities of fresh leaf tissue. Phytochem. Bull. 19: 11–15.

ERBAR, C. 1988. Early development patterns in flowers and their value for systematics. Pp. 7–23 in P. Leins, S. C. Tucker & P. K. Endress (editors), Aspects of Floral Development. J. Cramer, Berlin.

————. 1991. Sympetaly—a systematic character? Bot. Jahrb. Syst. 112: 417–451.

EYDE, R. 1988. Comprehending Cornus: puzzles and progress in the systematics of the dogwoods. Bot. Rev. (Lancaster) 54: 233–351.

FELSENSTEIN, J. 1978. Cases in which parsimony and compatibility methods will be positively misleading. Syst. Zool. 27: 401–410.

————. 1985. Confidence limits on phylogenies: an approach using the bootstrap. Evolution 39: 783–791.

FRIIS, E. M. 1985. Actinocalyx gen. nov., sympetalous angiosperm flowers from the Upper Cretaceous of southern Sweden. Rev. Paleobot. Palynol. 45: 171–183.

———— & W. L. CREPET. 1987. Time of appearance of floral features. Pp. 145–179 in E. M. Friis, W. G. Chaloner & P. R. Crane (editors), The Origins of Angiosperms and Their Biological Consequences. Cambridge Univ. Press, Cambridge.

GOLENBERG, E. M., D. E. GIANNASI, M. T. CLEGG, C. J. SMILEY, M. DURBIN, D. HENDERSON & G. ZURAWSKI. 1990. Chloroplast DNA sequence from a Miocene Magnolia species. Nature 344: 656–658.

GULOV, M. K., T. N. ULMASOV, K. A. ALIEV, V. M. ANDRIANOV & E. S. PIRUZIAN. 1990. Nucleotide sequence of the chloroplast rbcL gene from cotton Gossypium hirsutum. Nucleic Acids Res. 18: 187.

HAMBY, R. K. & E. A. ZIMMER. 1991. Ribosomal RNA as a phylogenetic tool in plant systematics. In P. Soltis, D. Soltis & J. J. Doyle (editors), Molecular Systematics of Plants. Routledge, Chapman & Hall, London.

HEGNAUER, R. 1977. The chemistry of the Compositae. Pp. 283–335 in V. H. Heywood, J. B. Harborne & B. L. Turner (editors), The Biology and Chemistry of the Compositae. Academic Press, London.

HENNIG, W. 1966. Phylogenetic Systematics. Univ. Illinois Press, Urbana, Illinois.

HUFFORD, L. D. 1992. Rosidae and their relationships to other nonmagnoliid dicotyledons: a phylogenetic analysis using morphological and chemical data. Ann. Missouri Bot. Gard. 79: 218–248.

JUSSIEU, A. L. DE. 1789. Genera Plantarum Secundum Ordines Naturales Disposita. Herissant & Barrois, Paris.

KIM, K.-J., R. K. JANSEN, R. S. WALLACE, H. J. MICHAELS & J. D. PALMER. 1992. Phylogenetic implications of rbcL sequence variation in the Asteraceae. Ann. Missouri Bot. Gard. 79: 428–445.

KLUGE, A. G. & J. S. FARRIS. 1969. Quantitative phyletics and the evolution of Anurans. Syst. Zool. 18: 1–32.

LES, D. H., D. K. GARVIN & C. F. WIMPEE. 1991. Molecular evolutionary history of ancient aquatic angiosperms. Proc. Natl. Acad. U.S.A. 88: 10119–10123.

LIN, C. M., Z. Q. LIU & S. D. KUNG. 1986. Nicotiana chloroplast genome: X. Correlation between the DNA sequences and the isoelectric focusing patterns of the LS of rubisco. Pl. Molec. Biol. 6: 81–87.

MARSHALL, C. R. 1990. The fossil record and estimating divergence times between lineages: maximum divergence times and the importance of reliable phylogenies. J. Molec. Evol. 30: 400–408.

MICHALOWSKI, C. B., H. J. BOHNERT, D. F. KLESSIG & J. O. BERRY. 1990. Nucleotide sequence of rbcL from Amaranthus hypochondriacus chloroplasts. Nucleic Acids Res. 18: 2187.

MITCHELL, L. G. & C. R. MERRIL. 1989. Affinity generation of single-stranded DNA for dideoxy sequencing following Polymerase Chain Reaction. Analytical Biochem. 178: 239–242.

MUELLER, J. 1981. Fossil pollen records of extant angiosperms. Bot. Rev. (Lancaster) 47: 1–142.

OLMSTEAD, R. G. & J. D. PALMER. 1992. A chloroplast DNA phylogeny of the Solanaceae: subfamilial relationships and character evolution. Ann. Missouri Bot. Gard. 79: 346–360.

————, R. K. JANSEN, H. J. MICHAELS, S. R. DOWNIE & J. D. PALMER. 1990. Chloroplast DNA and phylogenetic studies in the Asteridae. Pp. 119–134 in S. Kawano (editor), Biological Approaches and Evolutionary Trends in Plants. Academic Press, London.

PALMER, J. D. 1986. Isolation and structural analysis of chloroplast DNA. Methods Enzymol. 118: 167–186.

————, R. K. JANSEN, H. D. MICHAELS, M. W. CHASE & J. R. MANHART. 1988. Chloroplast DNA variation and plant phylogeny. Ann. Missouri Bot. Gard. 75: 1180–1206.

POLLARD, C. J. & K. S. AMUTI. 1981. Fructose oligosaccharides: possible markers of phylogenetic relationships among dicotyledonous plant families. Biochem. Syst. Ecol. 9: 68–78.

RITLAND, K. & M. T. CLEGG. 1987. Evolutionary analysis of plant DNA sequences. Amer. Naturalist 130: S74–S100.

RODMAN, J. E., M. OLIVER, R. NAKAMURA, J. McCLAMMER & A. BLEDSOE. 1984. A taxonomic analysis and revised classification of Centrospermae. Syst. Bot. 9: 297–323.

SAIKI, R. K., D. H. GELFAND, S. STOFFEL, S. SCHARF, R. H. HIGUCHI, G. T. HORN, K. B. MULLIS & H. A. ERLICH. 1988. Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase. Science 239: 487–491.

SANDERSON, M. J. & M. J. DONOGHUE. 1989. Patterns

Volume 79, Number 2
1992

Olmstead et al.   265
Monophyly of the Asteridae Inferred
from *rbc*L Sequences

of variation and levels of homoplasy. Evolution 43: 1781–1795.

SOLTIS, D. E., P. S. SOLTIS, M. T. CLEGG & M. DURBIN. 1990. *rbc*L sequence divergence and phylogenetic relationships in Saxifragaceae sensu lato. Proc. Natl. Acad. U.S.A. 87: 4640–4644.

SPORNE, K. R. 1969. The ovule as an indicator of evolutionary status in angiosperms. New Phytol. 68: 555–566.

———. 1975. The Morphology of Angiosperms: The Structure and Evolution of Flowering Plants. St. Martin's Press, New York.

STEBBINS, G. L. 1974. Flowering Plants. Evolution Above the Species Level. Belknap Press, Cambridge, Massachusetts.

SWOFFORD, D. L. 1989. PAUP: Phylogenetic Analysis Using Parsimony, version 3.0. Illinois Natural History Survey, Champaign, Illinois.

——— & G. J. OLSEN. 1990. Phylogeny reconstruction. Pp. 411–501 *in* D. M. Hillis & C. Moritz (editors), Molecular Systematics. Sinauer Associates, Sunderland, Massachusetts.

TAKHTAJAN, A. 1980. Outline of the classification of flowering plants (Magnoliophyta). Bot. Rev. (Lancaster) 46: 225–359.

———. 1987. Systema Magnoliophytorum. Officina Editoria Nauka, Leningrad. [In Russian.]

THORNE, R. F. 1983. Proposed new realignments in the angiosperms. Nordic J. Bot. 3: 85–117.

WAGENITZ, G. 1977. New aspects of the systematics of Asteridae. Pl. Syst. Evol. Suppl. 1: 375–395.

———. 1992. The Asteridae: evolution of a concept and its present status. Ann. Missouri Bot. Gard. 79: 209–217.

ZURAWSKI, G. & M. T. CLEGG. 1987. Evolution of higher-plant chloroplast DNA-encoded genes: implications for structure–function and phylogenetic studies. Ann. Rev. Pl. Physiol. 38: 391–418.

———, B. PERROT, W. BOTTOMLEY & P. R. WHITFELD. 1981. The structure of the gene for the large subunit of ribulose-1,5-bisphosphate carboxylase from spinach chloroplast DNA. Nucleic Acids Res. 14: 3251–3270.